



# Implications of adopting the Global Lungs Initiative 2012 all-age reference equations for spirometry

Philip H. Quanjer<sup>1</sup>, Danny J. Brazzale<sup>2</sup>, Piotr W. Boros<sup>3</sup> and Jeffrey J. Pretto<sup>4</sup>

**Affiliations:** <sup>1</sup>Dept of Pulmonary Diseases and Dept of Paediatrics, Erasmus Medical Centre, Erasmus University, Rotterdam, The Netherlands. <sup>2</sup>Dept of Respiratory and Sleep Medicine and Institute for Breathing and Sleep, Austin Hospital, Heidelberg, and <sup>4</sup>Dept of Respiratory and Sleep Medicine, John Hunter Hospital, School of Medicine and Public Health, University of Newcastle, Newcastle, Australia. <sup>3</sup>Lung Function Laboratory, National Tuberculosis and Lung Diseases Research Institute, Warsaw, Poland.

**Correspondence:** P.H. Quanjer, Kervel 19, 7443 GT Nijverdal, The Netherlands. E-mail: pquanjer@ziggo.nl

**ABSTRACT** The aim of this study was to determine the diagnostic and interpretative consequences of adopting the Global Lungs Initiative (GLI) 2012 spirometric prediction equations.

We assessed spirometric records from 17 572 subjects (49.5% females), aged 18–85 years, from hospitals in Australia and Poland.

We calculated predicted forced expiratory volume in 1 s (FEV<sub>1</sub>), forced expiratory volume (FVC), FEV<sub>1</sub>/FVC and lower limits of normal (LLN) using European Community for Steel and Coal (ECSC), National Health and Nutrition Examination Survey (NHANES) III and GLI 2012 equations. Obstruction was defined as FEV<sub>1</sub>/FVC < LLN and a restrictive pattern as FEV<sub>1</sub>/FVC > LLN and FVC < LLN. Global Initiative for Chronic Obstructive Lung Disease (GOLD) stage 2 and higher was defined as FEV<sub>1</sub>/FVC < 0.70 and FEV<sub>1</sub> < 80% pred.

GLI 2012 equations produce similar predicted values for FEV<sub>1</sub> and FVC compared with NHANES, but produce larger values than ECSC. Differences in the LLN lead to an important increase in the prevalence rate of a low FVC compared to ECSC, and a significant decrease compared to NHANES prediction equations. Adopting GLI 2012 equations has small effects on the prevalence rate of airway obstruction. GOLD stages 2–4 lead to >20% underdiagnosis of airway obstruction up to the age of 55 years and to 16–23% overdiagnosis in older subjects. GLI 2012 equations increase the prevalence of a “restrictive spirometric pattern” compared to ECSC but decrease it compared to NHANES.



@ERSpublications

Adopting the Global Lung Initiative 2012 prediction equations will have small effects on spirometric rates of detection <http://ow.ly/nrZxe>

---

For editorial comments see page 901.

Received: Dec 03 2012 | Accepted after revision: March 06 2013 | First published online: March 21 2013

Conflict of interest: None declared.

Copyright ©ERS 2013

## Introduction

Spirometry is the mainstay of lung function assessment, particularly in the diagnosis of obstructive lung disease. The diagnostic process entails comparing test results with expected values, which may be either previous measurements made in the same patient, or external reference values obtained from a disease-free dataset. Most physiological markers used for clinical purposes (such as serum concentrations of a substance) vary within narrow limits in healthy individuals because they are the outcome of a controlled system, and are independent of sex and ethnic group. However, spirometric indices and other measures of lung function vary with age, sex, standing height and ethnic group. This has led to a plethora of studies, primarily in Caucasians, documenting normal pulmonary function. Most studies include relatively small groups with limited age ranges leading to appreciable differences in predicted values between them [1–6], and to important disjunctions when individuals move from one age range to the next [7]. Consequently, the interpretation of test results varies with the predicted values used [1–6], and is also complicated by the fact that the coefficient of variation of spirometric indices varies with age, so that the lower limit of normal (LLN) is age dependent [8].

Recently a European Respiratory Society (ERS) Task Force (Global Lung Function Initiative (GLI)) addressed the above problems by producing predicted values for spirometric indices for a variety of ethnic groups, valid across the 3–95 years age range [9]. These GLI data were based on tests in >74 000 healthy nonsmokers from all over the world. As the report has been endorsed by the American Thoracic Society (ATS), the Australian and New Zealand Society of Respiratory Science, the Asian Pacific Society for Respiriology, the Thoracic Society of Australia and New Zealand and the American College of Chest Physicians, these equations are likely to be widely adopted and will probably supersede previous prediction equations. Inevitably, this will affect the interpretation of test results [3–6].

The objective of the present study is to evaluate the effect of a transition from two commonly used sets of prediction equations [10, 11] to the GLI 2012 equations on the classification of test results from a clinical population. The classifications are focused on interpretation of “airway obstruction”, “restrictive pattern” and “mixed pattern”. Since the Global Initiative for Chronic Obstructive Lung Disease (GOLD) guidelines [12] for diagnosing airway obstruction are still commonly used, we have also included these in the analysis.

## Methods

Consecutive clinical data were available from lung function laboratories in two tertiary hospitals in Australia: the John Hunter Hospital in New South Wales, the Austin Hospital in Victoria; and one in Poland: the National Tuberculosis and Lung Diseases Research Institute in Warsaw. The analysis was limited to Caucasians because there were very few non-Caucasian data. In addition, analysis was restricted to patients aged between 18 years (the lower age limit for one set of prediction equations [10]) and 85 years, because few data were available at higher ages. All spirometry data used were baseline or pre-bronchodilator status, and only test results meeting ATS/ERS criteria for acceptability and reproducibility were included [13]. Data were collected between August 2008 and June 2012 (John Hunter Hospital), January 2011 and May 2012 (Austin Hospital) and April 2009 and June 2012 (National Tuberculosis and Lung Diseases Research Institute). This study is a retrospective analysis of routinely obtained data which had been de-identified, obviating the need for approval from local ethics committees.

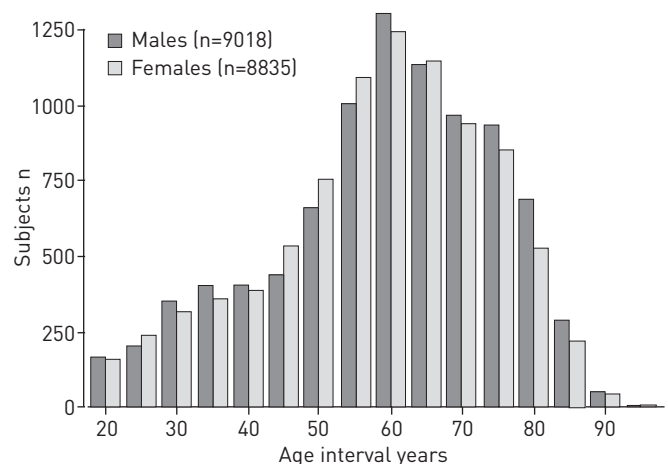


FIGURE 1 Age distribution of males and females in the clinical dataset. Each age group spans 5 years, thus age group 50 covers 47.51–52.50 years.

TABLE 1 Predicted values of spirometric indices in males (n=8870) and females (n=8702) aged 18–85 years, using the three predicted value studies [9–11]

	GLI 2012	ECSC/ERS	NHANES III
<b>Males</b>			
FEV <sub>1</sub> L	3.45 ± 0.69	3.28 ± 0.64	3.42 ± 0.72
FVC L	4.41 ± 0.78	4.14 ± 0.69	4.49 ± 0.77
FEV <sub>1</sub> /FVC	0.78 ± 0.02	0.77 ± 0.03	0.76 ± 0.03
<b>Females</b>			
FEV <sub>1</sub> L	2.49 ± 0.51	2.30 ± 0.54	2.50 ± 0.53
FVC L	3.12 ± 0.58	2.72 ± 0.58	3.19 ± 0.57
FEV <sub>1</sub> /FVC	0.80 ± 0.02	0.78 ± 0.03	0.79 ± 0.03

Data are presented as mean ± SD. GLI: Global Lung Initiative; ECSC: European Community for Steel and Coal; ERS: European Respiratory Society; NHANES: National Health and Nutrition Examination Survey; FEV<sub>1</sub>: forced expiratory volume in 1 s; FVC: forced vital capacity. All differences between predicted values are highly significant ( $p < 0.0001$ ).

Mean predicted values and Z-scores were derived using prediction equations from the European Community for Steel and Coal (ECSC)/ERS [10], National Health and Nutrition Examination Survey (NHANES) III [11] and GLI 2012 [9]. Z-scores according to the ECSC/ERS were calculated by (measured – predicted)/standard deviation of the predicted value. The NHANES III predicted values provide the mean and the LLN. As the LLN is defined as the fifth percentile, *i.e.* 1.6445 times the standard deviation below the predicted value, the standard deviation of the predicted values was calculated as (predicted – LLN)/1.6445. The GLI 2012 calculations [13] utilise mean predicted value, variation and skewness of the data to provide Z-scores and LLN specifically calculated for each individual. Airway obstruction was diagnosed if forced expiratory volume in 1 s (FEV<sub>1</sub>)/forced vital capacity (FVC) was <LLN, *i.e.* the Z-score was <-1.6445 [9–10, 14].

We limited the GOLD guidelines [12] analysis to stage 2 and higher, since GOLD stage 1 has been shown to be unrelated to respiratory disease [15–20]. GOLD-defined airway obstruction was considered to be present if FEV<sub>1</sub>/FVC <0.70 and FEV<sub>1</sub> <80% predicted.

A “restrictive pattern” in spirometry was diagnosed if FEV<sub>1</sub>/FVC was ≥LLN but the FVC <LLN. A “mixed pattern” was diagnosed if both FEV<sub>1</sub>/FVC and FVC were <LLN. In keeping with common pulmonary laboratory practice, predicted values and their LLN were calculated for the full age range, even though this requires extrapolation beyond the recommended age range for ECSC/ERS (70 years) and for NHANES III (80 years). Data analysis was performed using the statistical software R (version 2.15.1; R Foundation, [www.r-project.org](http://www.r-project.org)). Differences in prevalence rates according to the three prediction equations were tested using McNemar’s Chi-squared test and the test of differences between two dependent proportions. Differences in predicted values were analysed using the paired t-test.

## Results

27 patients were unusually short (<140 cm) and their data discarded, leaving 17 572 spirometry results for analysis: 10 783 from the National Tuberculosis and Lung Diseases Research Institute, 4129 from John Hunter Hospital and 2660 from Austin Hospital. The distribution by age is shown in [figure 1](#), with 49.5% being female. Predicted values ([table 1](#)) differed significantly ( $p < 0.0001$ ). The overall average predicted values from GLI 2012 and NHANES matched best, but NHANES predicted values and the LLN tended to be highest for FVC, and lowest for FEV<sub>1</sub>/FVC ([table 1](#) and [figs 2](#) and [3](#)). Mean predicted values for FEV<sub>1</sub> and FVC according to the ECSC/ERS were consistently smaller by 220–470 mL ([table 1](#) and [fig. 2](#)). The FEV<sub>1</sub>/FVC ratios predicted by ECSC/ERS and NHANES differed little, but were lower than those predicted by GLI 2012, particularly in those aged >60 years ([fig. 3](#)). The latter is mostly accounted for by the fact that the GLI 2012 data incorporates the inverse relationship between FEV<sub>1</sub>/FVC ratio and height, whereas this covariate is not included in the ECSC/ERS and NHANES prediction equations.

[Figure 4](#) illustrates the relative rates of airway obstruction (*i.e.* FEV<sub>1</sub>/FVC < LLN) for the three normal-value datasets evaluated and for the GOLD guidelines. Overall, the ECSC/ERS normal values generally resulted in the lowest rates of obstruction (31.1% for males and 24.7% for females), and GLI 2012 resulted in the highest rates (34.5% for males and 27.9% for females), with NHANES III generally being between the two (33.3% for males and 25.4% for females). While GOLD stages 2–4 rates of obstruction were similar on

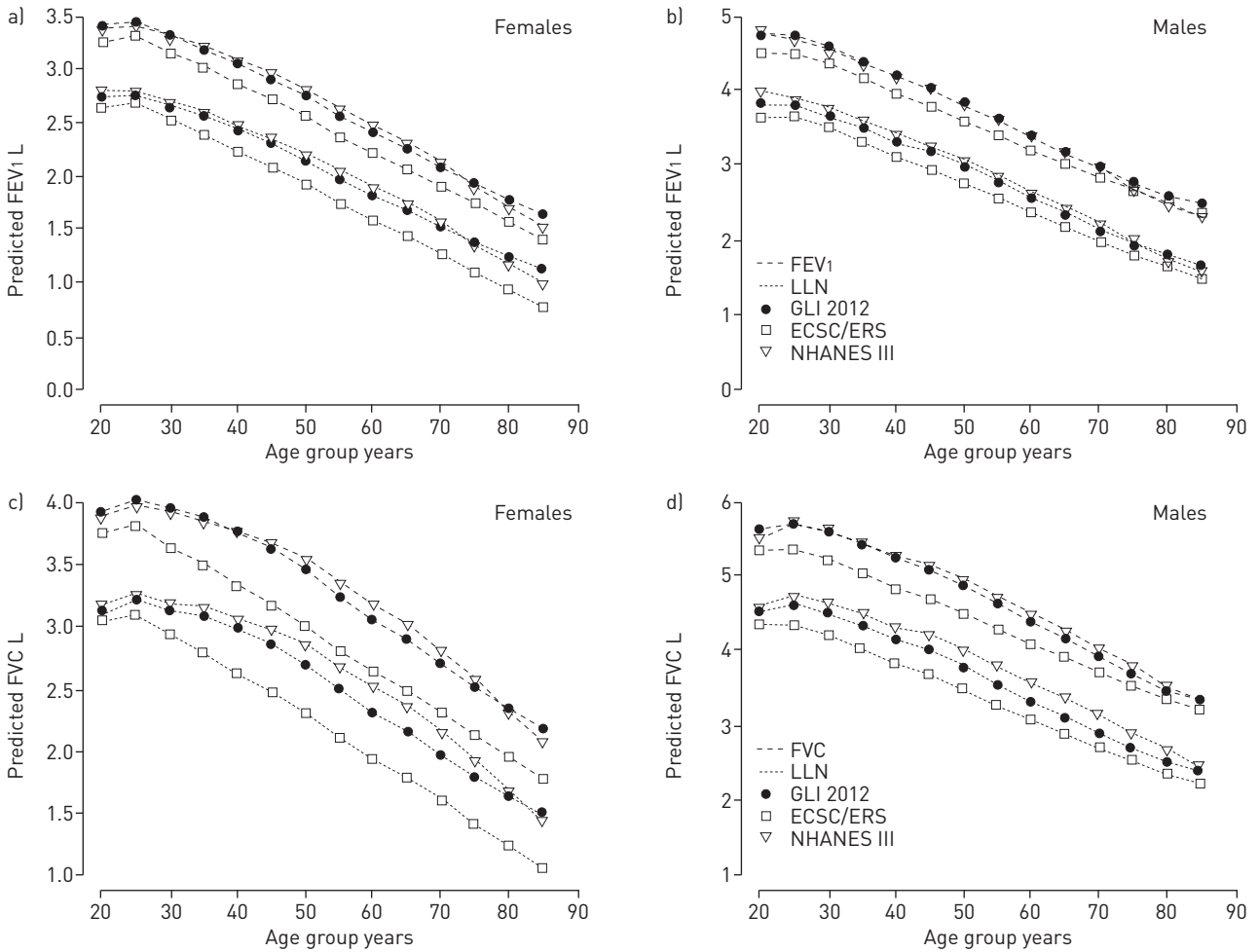


FIGURE 2 a) and b) predicted forced expiratory volume in 1 s (FEV<sub>1</sub>) and c) and d) forced vital capacity (FVC) and lower limits of normal (LLN) (fifth percentile) in the clinical dataset (n=17 572) by age group using prediction equations from the Global Lung Initiative (GLI) 2012 [9], European Community for Steel and Coal (ECSC)/European Respiratory Society (ERS) [10] and National Health and Nutrition Examination Survey (NHANES) III [11].

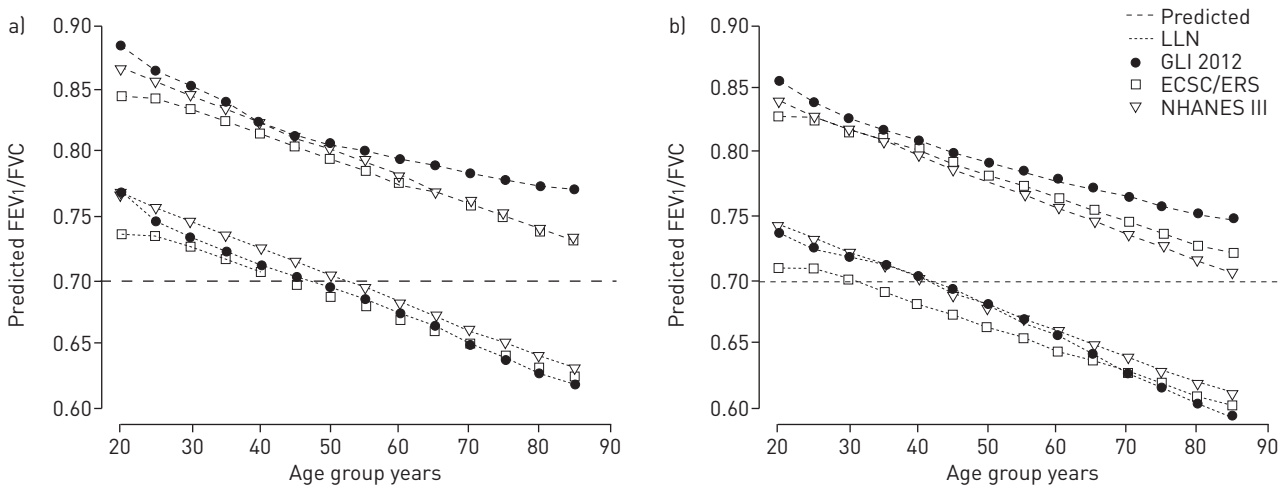


FIGURE 3 Predicted forced expiratory volume in 1 s (FEV<sub>1</sub>)/forced vital capacity (FVC) ratio and lower limits of normal (LLN) for a) females and b) males in the clinical dataset (n=17 572) by age group using prediction equations from Global Lung Initiative (GLI) 2012 [9], European Community for Steel and Coal (ECSC)/European Respiratory Society (ERS) [10] and National Health and Nutrition Examination Survey (NHANES) III [11].

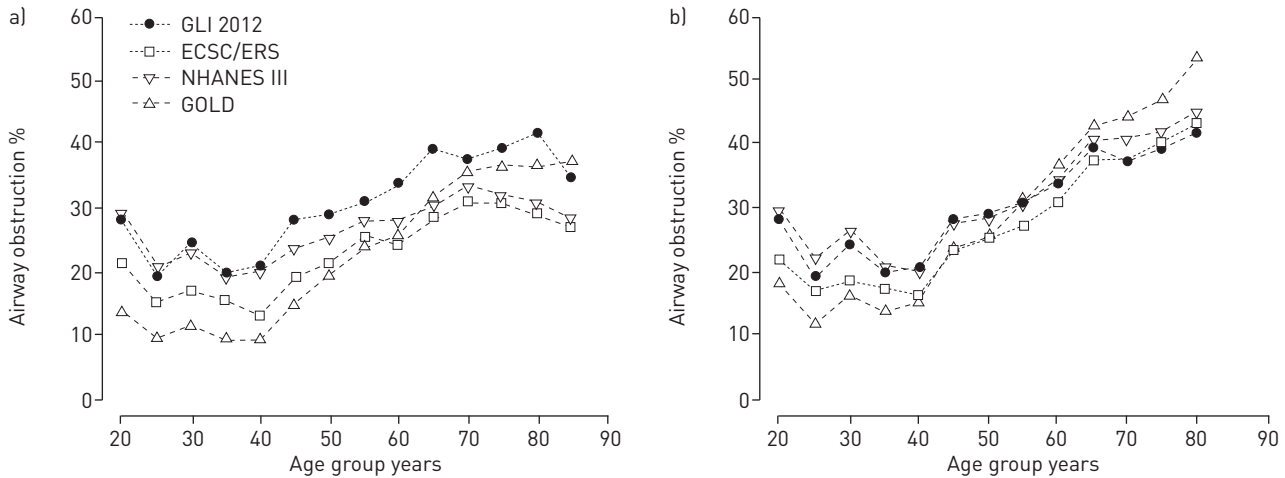


FIGURE 4 Percentage of subjects, arranged by age group, categorised as having airway obstruction defined as forced expiratory volume in 1 s (FEV<sub>1</sub>)/forced vital capacity (FVC)<lower limits of normal (LLN) (Global Lung Initiative (GLI) 2012, European Community for Steel and Coal (ECSC)/European Respiratory Society (ERS) and National Health and Nutrition Examination Survey (NHANES) III), or FEV<sub>1</sub>/FVC<0.70 and FEV<sub>1</sub> % predicted <80 (Global Initiative for Chronic Obstructive Lung Disease (GOLD) stages 2–4). a) Females; b) males.

average (35.1% and 25.4%), this averaging conceals the significantly lower rates in the young and the much higher rates in the older age group (>55–60 years) evident in figure 4.

Table 2 provides a breakdown of the level of FEV<sub>1</sub>/FVC and of FVC relative to their LLN. With the exception of airway obstruction in males using NHANES III equations (confidence interval of differences 0.051–0.017), all prevalence rates for ECSC/ERS and NHANES III differed significantly from those for GLI 2012 (maximum McNemar test p<0.019). The overall average rates of a low FVC (*i.e.* FVC<LLN irrespective of FEV<sub>1</sub>/FVC) were 20.9% for males and 12.8% for females using ECSC/ERS, 37.9% and 35.5%, respectively, for NHANES, and 29.1% and 27.8%, respectively, using GLI 2012; remarkable differences which reflect different levels of the LLN (fig. 2). Of these, 40.4–41.4% of males and 35.1–37.0% of females would be categorised as having a mixed pattern, *i.e.* both FEV<sub>1</sub>/FVC and FVC<LLN. Focusing on those with a purely restrictive-type pattern (*i.e.* FVC<LLN and FEV<sub>1</sub>/FVC>LLN), the overall average rates were 12.3% for males and 8.1% for females using ECSC/ERS, 22.6% and 22.8%, respectively, using NHANES III and 17.1% and 18.1%, respectively, using GLI 2012 (fig. 5).

TABLE 2 Subdivision of subjects according to the level of forced expiratory volume in 1 s (FEV<sub>1</sub>)/forced vital capacity (FVC) and FVC relative to their lower limits of normal (LLN)

	Male <sup>#</sup>			Female <sup>†</sup>		
	GLI 2012	ECSC/ERS	NHANES III	GLI 2012	ECSC/ERS	NHANES III
<b>Obstruction</b>	33.3	31.1	34.5	25.4	24.7	27.9
<b>Obstruction and FEV<sub>1</sub>&lt;LLN</b>	26.4	22.7	27.6	20.4	15.7	22.1
<b>Obstruction and FVC&gt;LLN</b>	21.2	22.5	19.2	15.6	20.0	15.1
<b>FVC&lt;LLN</b>	29.1	20.9	37.9	27.8	12.8	35.5
<b>Restrictive pattern</b>	17.1	12.3	22.6	18.1	8.1	22.8
<b>Mixed pattern</b>	12.1	8.6	15.3	9.7	4.7	12.7

Data are presented as %. All prevalence rates differ significantly (p<0.019) from Global Lung Initiative (GLI) 2012, except for overall obstruction between GLI 2012 and NHANES (National Health and Nutrition Examination Survey) III in males. Obstruction: FEV<sub>1</sub>/FVC<LLN; restrictive pattern: FVC<LLN and FEV<sub>1</sub>/FVC>LLN; mixed pattern: FEV<sub>1</sub>/FVC<LLN and FVC<LLN; ECSC: European Community for Steel and Coal; ERS: European Respiratory Society. #: n=8870; †: n=8707.

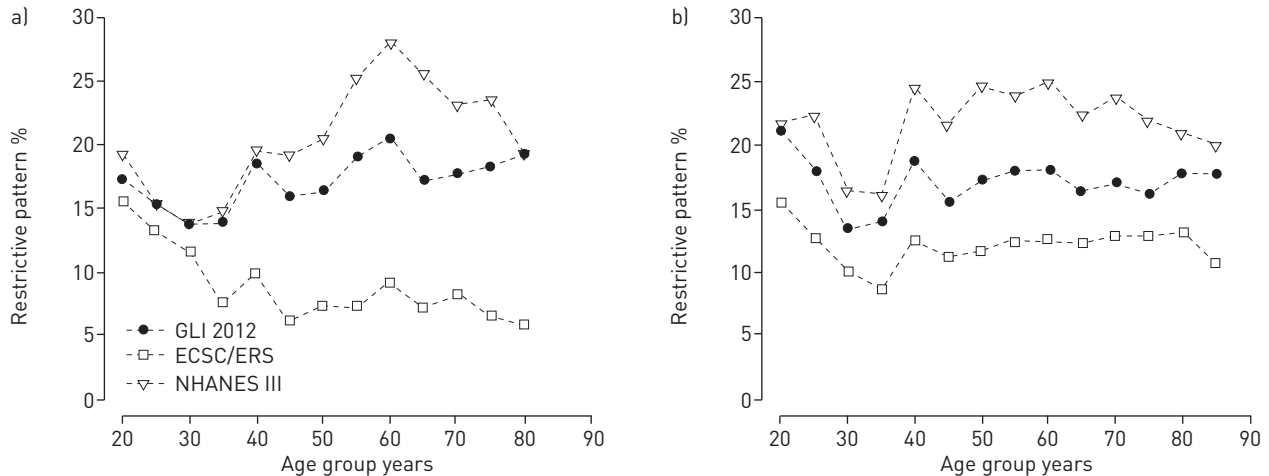


FIGURE 5 Percentage of subjects, arranged by age group, categorised as having a restrictive pattern, *i.e.* forced expiratory volume in 1 s /forced vital capacity (FVC)>lower limits of normal (LLN) and FVC<LLN in a) females and b) males. GLI: Global Lung Initiative; ECSC: European Community for Steel and Coal; ERS: European Respiratory Society; NHANES: National Health and Nutrition Examination Survey.

### Discussion

This study is based on the premise that measurements that fall below the upper 95% distribution from a large sample of a healthy population of nonsmokers are, in principle, indicative of a disease process. We regarded the GLI 2012 predicted values [9] as the reference standard because they are based on a large population sample from all over the world (25 827 and 31 568 Caucasian males and females, respectively), and because the modelling used in their development led to normally distributed residuals. This allows lung function to be expressed as a Z-score, adjusted for height, age, sex and ethnic group [9]. Z-scores indicate how much a measurement differs from the mean predicted value, with only 5% of healthy subjects having a Z-score  $\leq -1.6445$  (fifth percentile). As is common in clinical measurement, we adopted the fifth percentile as the LLN, realising that this cutoff value will result in a 1 in 20 false-positive rate. This is unlikely to be problematic given that our analysis was conducted on data from patients referred to pulmonary laboratories of tertiary hospitals, so that there is often associated evidence of respiratory disease. Clinical diagnoses require more than pulmonary function test results and incorporate factors such as patient symptoms, relevant history and clinical signs. As such, we have limited our analysis to the identification of three nonspecific syndromes rather than to detection of specific disease states. These are airway obstruction, a restrictive pattern and a mixed pattern, which commonly form the basis for interpretative strategies in clinical respiratory medicine [14].

Changing predicted values creates a number of potential problems for laboratories and clinicians, and we consider it important to identify how clinical interpretation of test results might alter on adoption of the new GLI 2012 reference data. Issues such as effects on presence and severity of disease, as well as the interpretation of serial observations are raised. Interpretation of lung function is still widely based on results expressed as % pred, despite the fact that this leads to significant age, height, sex and ethnic group-related bias [8–10, 21–25]. In addition, the interpretation of serial measurements may be especially confounded if they are evaluated on the basis of % pred without awareness of the fact that there has been a change in reference equations [26].

We have found that predicted values for FEV<sub>1</sub> and FVC from GLI 2012 did not differ grossly (table 1) compared to the NHANES III study, a high quality study [11]. Although NHANES study data do comprise some of the GLI 2012 data and so some congruence may be expected, they contributed <4% of the total dataset, which is insufficient to explain this agreement. Transition from the predicted values of NHANES III [11], which had been recommended for use in USA [27], to GLI 2012 equations will therefore have a limited impact on the predicted FEV<sub>1</sub> and FVC, except in more elderly subjects. We suspect that this finding in the elderly may result from extrapolation beyond the valid age range (for both NHANES III and ECSC/ERS), leading to a steeper fall in the LLN in elderly subjects than that of the GLI 2012 data (fig. 2). Differences in the LLN for FVC (fig. 2) led to up to a threefold difference in the prevalence ratio of a low FVC (table 2). Predicted values from the ECSC/ERS [10], which are widely used in Europe and elsewhere, are appreciably lower across all ages; this is not a new finding [6, 28].



There are more pronounced differences in the predicted values for FEV<sub>1</sub>/FVC (fig. 3); this is because ECSC/ERS and NHANES III, unlike GLI 2012, do not take into account that the FEV<sub>1</sub>/FVC ratio falls with increasing standing height (which is lower in elderly subjects). The predicted FEV<sub>1</sub>/FVC ratio tends to be highest for GLI 2012 (fig. 3). The ECSC/ERS equations lead to lower rates of diagnosis of airway obstruction up to ~60 years of age, much less so in females than in males (fig. 4).

Inevitably there will be overlap in the distribution of spirometric test results between healthy subjects and those with disease. Therefore, adoption of the LLN derived from healthy subjects for clinical purposes is open to criticism. The GOLD group advocates the use of FEV<sub>1</sub>/FVC <0.70 as indicating pathological airflow limitation, acknowledging that using the LLN would lower the high rate of false-negative and false-positive results [12]. While the LLN methodology is based on an appropriate statistical approach, the clinical validity of either approach hinges on evidence of respiratory disease being present in those subjects who meet the respective inclusion criteria. It has been shown that GOLD stage 1 is not associated with premature death [15, 17, 19, 29], abnormal decline in FEV<sub>1</sub>, respiratory symptoms, respiratory care nor quality of life [16]. Conversely, GOLD stage 1 with FEV<sub>1</sub>/FVC < LLN is associated with increased risk of death [16], confirming the finding by MANNINO *et al.* [30] that the adjusted hazard ratio for premature death in subjects with GOLD stage 1 was only significant if FEV<sub>1</sub>/FVC was < LLN. The use of LLN for FEV<sub>1</sub>/FVC and FEV<sub>1</sub> identifies persons with increased risk of death and prevalence of respiratory symptoms [18]. All the evidence therefore suggests that using the LLN for FEV<sub>1</sub>/FVC to identify subjects with pathological airflow limitation is clinically valid, unlike the fixed ratio advocated by the GOLD group [12].

The GOLD classification for stage 2 and higher is associated with a pronounced age-related bias. Our data shows that in patients ≤55 years of age there is a systematically low rate of diagnosis, so that 24% of cases of abnormal airflow obstruction are missed. In older subjects there is 23% overdiagnosis in males, and 16% in females. This corroborates an earlier report of a misclassification rate of ~20% using GOLD criteria [21]. The percentage misclassification would have been much higher had we also included GOLD stage 1 in our analysis [3]. Our findings, and those of MILLER *et al.* [21], underline that even GOLD stage 2 classification has serious shortcomings which invalidate its clinical use.

In 1991 the ATS [22] suggested that a restrictive ventilatory defect, *i.e.* a condition with a reduced total lung capacity, may be inferred when the vital capacity is reduced and FEV<sub>1</sub>/FVC is normal or increased. Since then a “spirometric restrictive pattern” is increasingly mentioned in the literature and widely used in clinical practice. Our analysis shows that the rates of detecting this pattern vary substantially with normal dataset used, with the lowest prevalence generally with the ECSC/ERS equations (fig. 5). As such, this restrictive pattern will be detected more frequently when laboratories switch from using ECSC/ERS, as previously described [6]. This may increase the referral rate for assessment of total lung capacity to confirm the restrictive defect, thereby increasing medical expenditure. The reality is that the veracity of the 1991 ATS approach for interpreting a restrictive pattern from spirometry is questionable [31–33] and even worse when there is concomitant airways obstruction [14, 33–35]. The restrictive pattern and mixed pattern are elusive syndromes with a very high negative predictive value but a poor positive predictive value for reduced total lung capacity. In a general population, where a reduced total lung capacity is rare, the positive predictive value of this test result will be very small. There is therefore little to be said for measurement of the total lung capacity if a restrictive pattern or mixed pattern is diagnosed on the basis of spirometry alone, unless there is *a priori* clinical evidence of a condition that might cause lung restriction, such as interstitial lung disease [36]. The underlying principle should be to treat the patient, not the numbers [37].

A clinical diagnosis is based on clinical signs, symptoms and objective measures, among which pulmonary function test results play a prominent role. Assessing whether there is evidence of airway obstruction, or a restrictive pattern, forms part of the clinical routine, and helps in assessing disease severity, prognosis and interventions of best potential utility. Against that background this study provides useful information to what extent clinical decision making might be affected by the transition to the GLI 2012 prediction equations. This study is based on a hospital population with considerably higher disease prevalence than in general practice or occupational medicine, and so the prevalence of airways obstruction and restrictive spirometric patterns would be expected to be higher.

The strength of the GLI 2012 equations is that they have been derived from a large, representative population sample [38], are valid over a wide age range without any discontinuity between age groups, and apply to different ethnic groups. This study is limited to Caucasians in the 18–95-year age range; hence complementary studies are required to evaluate the consequences of the transition to GLI 2012 equations in paediatrics and in non-Caucasians. A potential limitation of this study is that the analysis uses pre-bronchodilator values, whereas the GOLD classification utilises post-bronchodilator values [12]. The GOLD statement focuses on chronic obstructive pulmonary disease (COPD) which, being defined as irreversible airway obstruction, requires reversibility to be excluded to confirm the diagnosis. This study, however, is

not about COPD in isolation and is more wide-ranging across all disease types. Furthermore, the appropriate reference value to use for post-bronchodilator spirometry is one obtained in healthy subjects who were administered the same drug. Bronchodilator use increases the FEV<sub>1</sub>/FVC ratio on average by 0.013–0.029 in healthy individuals, depending on age and sex [39]. It is therefore likely that using post-bronchodilator data and post-bronchodilator reference values in our analysis would have resulted in similar findings.

In conclusion, the transition from NHANES III to GLI 2012 spirometry reference data will lead to limited changes in the predicted values of FEV<sub>1</sub> and FVC for adults. In contrast, users of ECSC/ERS prediction equations should be aware that predicted values for FEV<sub>1</sub> and FVC will increase by a few hundred millilitres on transition to GLI 2012 data. This will lead to a substantial increase in the prevalence rate of low FVC when changing from ECSC/ERS, or a moderate decrease if changing from NHANES III. There are substantial differences in the predicted levels of FEV<sub>1</sub>/FVC ratio but not in their LLN, so that adopting the GLI 2012 prediction equations will have small effects on the rates of detection of obstructive ventilatory defects. Our analysis confirms previous findings in clinical populations that GOLD definitions for stage 2 and higher lead to >20% underdiagnosis of airway obstruction in those aged ≤55 years, and to 16–23% overdiagnosis in older subjects.

## References

- 1 Quanjer PH. Standardized lung function testing. Report Working Party Standardization of Lung Function Tests, European Community for Steel and Coal. *Bull Eur Physiopath Respir* 1983; 19: Suppl. 5, 45–51.
- 2 Quanjer PH, Stocks J, Polgar G, *et al.* Compilation of reference values for lung function measurements in children. *Eur Respir J* 1989; 2: Suppl. 4, 184S–261S.
- 3 Swanney MP, Ruppel G, Enright PL, *et al.* Using the lower limit of normal for the FEV<sub>1</sub>/FVC ratio reduces the misclassification of airway obstruction. *Thorax* 2008; 63: 1046–1051.
- 4 Sood A, Dawson BK, Henkle JQ, *et al.* Effect of change of reference standards to NHANES III on interpretation of spirometric “abnormality”. *Int J Chron Obstruct Pulm Dis* 2007; 2: 361–367.
- 5 Collen J, Greenburg D, Holley A, *et al.* Discordance in spirometric interpretations using three commonly used reference equations vs National Health and Nutrition Examination Study III. *Chest* 2008; 134: 1009–1016.
- 6 Brazzale DJ, Upward AL, Pretto JJ. Effects of changing reference values and definition of the normal range on interpretation of spirometry. *Respirology* 2010; 15: 1098–1103.
- 7 Quanjer PH, Hall GL, Stanojevic S, *et al.* Age- and height-based prediction bias in spirometry reference equations. *Eur Respir J* 2012; 40: 190–197.
- 8 Stanojevic S, Wade A, Stocks J, *et al.* Reference ranges for spirometry across all ages: a new approach. *Am J Respir Crit Care Med* 2008; 177: 253–260.
- 9 Quanjer PH, Stanojevic S, Cole TJ, *et al.* Multi-ethnic reference values for spirometry for the 3–95-yr age range: the global lung function 2012 equations. *Eur Respir J* 2012; 40: 1324–1343.
- 10 Quanjer PH, Tammeling GJ, Cotes JE, *et al.* Lung volumes and forced ventilatory flows. Report Working Party Standardization of Lung Function Tests, European Community for Steel and Coal. Official Statement of the European Respiratory Society. *Eur Respir J* 1993; 6: Suppl. 16, 5S–40S.
- 11 Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. *Am J Respir Crit Care Med* 1999; 159: 179–187.
- 12 Global Initiative for Chronic Obstructive Lung Disease. Global Strategy for Diagnosis, Management and Prevention of Chronic Obstructive Lung Disease. [www.goldcopd.org/guidelines-global-strategy-for-diagnosis-management.html](http://www.goldcopd.org/guidelines-global-strategy-for-diagnosis-management.html) Date last accessed: September 29, 2011. Date last updated: February 2013.
- 13 Miller MR, Hankinson J, Brusasco V, *et al.* Standardisation of spirometry. *Eur Respir J* 2005; 26: 319–338.
- 14 Pellegrino R, Viegi G, Brusasco V, *et al.* Interpretative strategies for lung function tests. *Eur Respir J* 2005; 26: 948–968.
- 15 Ekberg-Aronsson M, Pehrsson K, Nilsson JA, *et al.* Mortality in GOLD stages of COPD and its dependence on symptoms of chronic bronchitis. *Respir Res* 2005; 6: 98.
- 16 Bridevaux PO, Gerbase MW, Probst-Hensch NM, *et al.* Long-term decline in lung function, utilisation of care and quality of life in modified GOLD stage 1 COPD. *Thorax* 2008; 63: 768–774.
- 17 Vaz Fragoso CA, Concato J, McAvay G, *et al.* Chronic obstructive pulmonary disease in older persons: a comparison of two spirometric definitions. *Respir Med* 2010; 104: 1189–1196.
- 18 Vaz Fragoso CA, Concato J, McAvay G, *et al.* The ratio of FEV<sub>1</sub> to FVC as a basis for establishing chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2010; 181: 446–451.
- 19 Pedone C, Scarlata S, Sorino C, *et al.* Does mild COPD affect prognosis in the elderly? *BMC Pulm Med* 2010; 10: 35.
- 20 Vaz Fragoso CA, Gill TM, McAvay G, *et al.* Use of lambda-mu-sigma-derived Z score for evaluating respiratory impairment in middle-aged persons. *Respir Care* 2011; 56: 1771–1777.
- 21 Miller MR, Quanjer PH, Swanney MP, *et al.* Interpreting lung function data using 80% predicted and fixed thresholds misclassifies more than 20% of patients. *Chest* 2011; 139: 52–59.
- 22 Lung function testing: selection of reference values and interpretative strategies. American Thoracic Society. *Am Rev Respir Dis* 1991; 144: 1202–1218.
- 23 Sobol BJ. Assessment of ventilatory abnormality in the asymptomatic subject: an exercise in futility. *Thorax* 1966; 2: 445–449.
- 24 Sobol BJ, Sobol PG. Per cent of predicted as the limit of normal in pulmonary function testing: a statistically valid approach. *Thorax* 1979; 34: 1–3.
- 25 Miller MR, Pincock AC. Predicted values: how should we use them? *Thorax* 1988; 43: 265–267.
- 26 Kirkby J, Aurora P, Spencer H, *et al.* Stitching and switching: the impact of discontinuous lung function reference equations. *Eur Respir J* 2012; 39: 1256–1257.



- 27 Global Lung Function Initiative. GLI 2012 Data Conversion Software. [www.lungfunction.org/files/InstallGLI2012\\_DataConversion.EXE](http://www.lungfunction.org/files/InstallGLI2012_DataConversion.EXE) Date last accessed: March 1, 2013. Date last updated: May 10, 2013.
- 28 Roca J, Burgos F, Sunyer J, *et al.* Reference values for forced spirometry. Group of the European Community Respiratory Health Survey. *Eur Respir J* 1998; 11: 1354–1362.
- 29 Mannino DM, Doherty DE, Buist SA. Global Initiative on Obstructive Lung Disease (GOLD) classification of lung disease and mortality: findings from the Atherosclerosis Risk in Communities (ARIC) study. *Respir Med* 2006; 100: 115–122.
- 30 Mannino DM, Buist SA, Vollmer WM. Chronic obstructive pulmonary disease in the older adult: what defines abnormal lung function? *Thorax* 2007; 62: 237–241.
- 31 Aaron SD, Dales RE, Cardinal P. How accurate is spirometry at predicting restrictive pulmonary impairment? *Chest* 1999; 115: 869–873.
- 32 Gladly CA, Aaron SD, Lunau ML, *et al.* A spirometry-based algorithm to direct lung function testing in the pulmonary function laboratory. *Chest* 2003; 123: 1939–1946.
- 33 Swanney MP, Beckert LE, Frampton CM, *et al.* Validity of the American Thoracic Society and other spirometric algorithms using FVC and forced expiratory volume at 6 s for predicting a reduced total lung capacity. *Chest* 2004; 126: 1861–1866.
- 34 Boros P, Franczuk M, Wesolowski S. Zmiany “mieszane” – weryfikacja rozpoznania typu zaburzeń wentylacyjnych [“Mixed” changes in spirometry-verification of the pattern of lung function impairment]. *Pneumonol Alergol Pol* 2003; 71: 527–532.
- 35 Dykstra BJ, Scanlon PD, Kester MM, *et al.* Lung volumes in 4774 patients with obstructive lung disease. *Chest* 1999; 115: 68–74.
- 36 Boros PW, Franczuk M, Wesolowski S. Value of spirometry in detecting volume restriction in interstitial lung disease patients: spirometry in interstitial lung diseases. *Respiration* 2004; 71: 374–379.
- 37 Saunders KB. *Clinical Physiology of the Lung*. Oxford, Blackwell Science Publications, 1977.
- 38 Quanjer PH, Stocks J, Cole TJ, *et al.* Influence of secular trends and sample size on reference equations for lung function tests. *Eur Respir J* 2011; 37: 658–664.
- 39 Johannessen A, Lehmann S, Omenaas ER, *et al.* Post-bronchodilator spirometry reference values in adults and implications for disease management. *Am J Respir Crit Care Med* 2006; 173: 1316–1325.